

## University of Groningen

### Belief biased reasoning in anxiety disorders

Vroling, Maartje Sophie

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2011

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Vroling, M. S. (2011). *Belief biased reasoning in anxiety disorders*. s.n.

#### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

#### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

# Belief bias and the extinction of induced fear

M. S. Vroling & P. J. de Jong

Submitted for publication

## Abstract

Phobic individuals expect aversive UCSs when encountering phobic stimuli, while in reality such a sequential relationship is non-existent. It is unclear why these UCS expectancies do not extinguish in the face of disconfirming evidence. Extinction requires that people deduce the logical implication of corrective experiences that challenge the previously learned CS-UCS contingency. Therefore, a strong tendency to confirm a priori beliefs may be a relatively direct and powerful mechanism immunizing against refutation of once acquired UCS expectancies. The present study was designed to investigate whether individual differences in habitual (belief biased) reasoning may help explaining individual differences in the pattern of extinction. We tested whether relatively strong belief-confirming reasoning (belief bias) predicts delayed extinction of experimentally induced UCS expectancies. In a differential aversive conditioning paradigm, we used UCS-irrelevant (Exp. 1,  $N = 74$ ) and UCS-relevant (Exp. 2,  $N = 176$ ) pictorial stimuli as the CS<sup>+</sup> and electrical stimulation as the UCS. Belief bias was not or negatively related to extinction when a priori CS-UCS belongingness was absent (Exp. 1), whereas belief bias did predict delayed extinction of UCS expectancies when there was a high a priori CS-UCS belongingness (as is typically the case for phobic stimuli, Exp. 2). Together these findings indicate that enhanced belief bias may indeed play a role in the persistence of non-realistic anxiogenic UCS expectancies, thereby contributing to the development and persistence of anxiety disorders. It also points to the relevance of reasoning tendencies in the search for predictors of delayed extinction of UCS expectancies.

## Introduction

Early learning theory accounts of anxiety disorders have claimed that anxiety disorders are caused by aversive learning experiences (either directly or vicariously; e.g., Öst & Hugdahl, 1981; Watson & Rayner, 1920). Contemporary learning theories emphasize the additional importance of the influence of prior experiences, post event processing, on-site interpretations and individual variability in sensitivity for aversive learning experiences (Mineka & Zinbarg, 2006). Individual variability in the sensitivity for aversive and corrective experiences is assumed to be an important factor in the likelihood for developing an anxiety disorder. In line with this, recent conditioning research shows that people with panic disorder display delayed extinction of conditioned fear compared to healthy controls (Michael, Blechert, Vriends, Margraf, & Wilhelm, 2007). Yet, thus far little is known about what exactly constitutes these individual differences in conditionability (i.e., rate of extinction).

Successful extinction requires that people are sensitive to corrective experiences that challenge the previously learned CS-UCS contingency (e.g., this time the CS<sup>+</sup> was not followed by the UCS). If people have a (habitual) difficulty with incorporating disconfirming information, somehow acquired UCS expectancies will be more difficult to extinguish. Individual differences in extinction may thus be (partly) explained by individual differences in people's habitual tendency to neglect information that is inconsistent with these prior beliefs.

One reasoning tendency that directly relates to the incorporation of threat-disconfirming information (which is presented during extinction) is belief bias: Belief bias refers to the difficulty to distinguish what one believes from what is logically valid (e.g., Evans, Newstead, & Byrne, 1993). Belief bias is a common reasoning bias which serves to maintain strongly held beliefs: If people are confronted with evidence that goes against what they strongly believe, they tend to focus more on the believability of the information than on the logical validity (e.g., Evans, Over, & Manktelow, 1993). In line with the hypothesis that a relatively strong habitual belief bias may immunize against refutation of biased UCS expectancies, there is evidence that phobic individuals are characterized by a generally enhanced belief bias (de Jong, Weertman, Horselenberg, & van den Hout, 1997).

The interference of prior beliefs on logical reasoning performance is commonly measured using syllogisms. Previous research on belief biased reasoning in the context of anxiety (e.g., de Jong, Weertman et al., 1997; Vroling & de Jong, 2010b) relied on linear syllogisms: A linear syllogism consists of two premises and a conclusion. Participants need to assume that the premises are true, and need to determine whether or not the conclusion is logically valid.

Logically valid refers to the conclusion necessarily following from the premises. An example of a logically valid syllogism is:

Premise 1	A is larger than B
Premise 2	B is larger than C
<b>Conclusion</b>	<b>A is larger than C</b>

Given that these premises are true, it would be a violation of the rules of logic to conclude that ‘C is larger than A’. In order to measure belief bias, participants are asked to indicate whether or not syllogisms are logically valid, while both the logical validity and the believability of the conclusions are varied. Participants are instructed to ignore the believability of the syllogisms. An example of the four possible variations of a syllogism is presented in Table 3.1. When the content of the syllogisms is related to strongly held beliefs, participants show slower responses and/or more errors for syllogisms in which the logical validity and the believability do not match. This is known as the belief bias effect (e.g., Evans, Newstead et al., 1993).

Table 3.1  
*Example of a linear syllogism varying in believability of the conclusion and logical validity.*

	Believable conclusion	Unbelievable conclusion
valid	A mouse is smaller than a dog	An elephant is smaller than a dog
	A dog is smaller than an elephant	A dog is smaller than a mouse
	<b>A mouse is smaller than an elephant</b>	<b>An elephant is smaller than a mouse</b>
invalid	An elephant is smaller than a dog	A mouse is smaller than a dog
	A dog is smaller than a mouse	A dog is smaller than an elephant
	<b>A mouse is smaller than an elephant</b>	<b>An elephant is smaller than a mouse</b>

People with a relatively strong belief bias will likely tend to persist in their somehow acquired belief that the CS<sup>+</sup> is a predictor of the UCS and will thus show a difficulty to learn that a formerly threatening stimulus is now safe (viz. in the extinction phase the CS<sup>+</sup> is no longer followed by the UCS). The major aim of the present study was to test whether relatively strong belief bias is indeed associated with delayed extinction of experimentally induced UCS expectancies in the context of a differential aversive conditioning paradigm. As a secondary issue, we explored whether belief bias may also moderate the acquisition of differential UCS expectancies.

# Experiment 1

## Method

### Participants

Participants ( $N = 74$ , 23 male and 51 female) were undergraduate students in psychology (83.8%) and other faculties. Their mean age was 20.46 years ( $SD = 2.20$ ). The participating psychology students received course credits, the other students received a small financial reward. Students who were not fluent in Dutch, who suffered from dyslexia, or who had received training in logical reasoning were excluded from entering the study to avoid artificial noise in the belief bias data.

### Materials and apparatus

**Belief bias task.** Belief bias was measured using a computerized syllogistic reasoning task. Participants were asked to judge as quickly as possible whether the conclusion logically followed from the premises. The presented syllogisms varied in logical validity and in believability of the conclusion. A belief bias effect is found when participants are faster and/or make less mistakes when there is a match between the logical validity and the believability of a syllogism (i.e., the valid-believable and the invalid-unbelievable syllogisms) than when solving syllogisms for which logical validity and believability do not match (i.e., the valid-unbelievable and the invalid-believable syllogisms).

The belief bias task consisted of neutral syllogisms similar to the neutral syllogisms used in Vroling and de Jong (2010b). There were four different syllogisms, and each syllogism was varied as a function of believability and validity. Furthermore, each syllogism was presented in two orders to counter possible reading strategies ( $a > b$ ,  $b > c$ , therefore  $a > c$  and  $b > c$ ,  $a > b$ , therefore  $a > c$ ). The  $4$  (syllogisms)  $\times 4$  (types)  $\times 2$  (orders) = 32 resulting syllogisms were presented in two blocks<sup>5</sup>. The blocks were separated by a 30 s. break (cf. Vroling & de Jong, 2010b). The syllogisms were presented in a fixed random order, with the following restrictions: The topic should differ with every stimulus presentation, type of syllogism should differ after a maximum of two stimulus presentations, and order should differ after a maximum of three stimulus presentations. The outcome measures were reaction time (RT) and number of errors.

Errors and RTs reflect two different processes: Both belief bias measured in errors and measured in RTs indicate interference of believability in logical

---

<sup>5</sup> For pilot-purposes, we also included four threat related syllogisms, which we also varied as a function of believability and validity. Therefore the complete reasoning task consisted of 64 (two blocks of 32) syllogisms.

reasoning performance, yet people who make belief-based reasoning *mistakes* have more difficulty to distinguish between believability and logical validity than people who just need more time to come up with the correct answer. In a lab setting, participants scoring high on RT-based belief bias can be described as critical reasoners, since they correctly identify that their initial belief-based response-tendency is in fact incorrect.<sup>6</sup>

**Differential fear conditioning task.** Two intrinsically neutral stimuli were used throughout the fear conditioning task: a blue half circle, rotated 45° to the left and a blue half circle rotated 45° to the right. The stimuli were presented on a gray-shaded background, and were projected with a beamer on a white screen approximately 3 m in front of the subject. On-line probability estimates were measured with a rotary lever on a 0-100 scale (range 180°), positioned in front of the seated subject.

The aversive electrical tactile shock was administered on the middle and ring finger of the dominant hand by means of two Ag/AgCl electrodes (diameter 8 mm). To guarantee the safety of the participants, the shock was only administered after the current had been directed through an SHK1 isolation shocker (PsyLab) with a range of 0 to 5 mA.

**Fear conditioning questionnaire.** Participants were asked to indicate which of the two stimuli was (sometimes) followed by a shock, and had to rate how certain they were of their decision on a visual analogue scale (VAS, 0-100 mm ranging from ‘very uncertain’ to ‘very certain’).

### Procedure

After filling out an informed consent, participants started with the belief bias task. Participants were instructed to judge the validity of the syllogisms as quickly as possible by pressing a red ‘NO’ key on the left side or a green ‘YES’ key on the right side of an E-prime response box. Participants were given four practice items with feedback on the correctness of their response. Further explanation on the validity of the conclusion was given on the first two exemplars. Instructions were repeated after the 30 s break.

Each stimulus was preceded by a blank screen (500 ms) and a screen reading ‘pay attention!’ (1500 ms). Each stimulus disappeared as soon as a response was given, with a maximum of a 20 second delay before the response was coded ‘incorrect’. No feedback was given during the test-phase.

When participants had completed the belief bias task, they were seated in front of a projection screen for the differential fear conditioning task. Participants

---

<sup>6</sup> It should be noted that people showing high RT-based belief bias measured in the controlled environment of a lab will probably show belief-based reasoning *errors* in everyday life, since everyday life does not usually provide a single-task 20 s decision opportunity.

were informed that, for this part of the experiment, stimuli would be presented on the screen and that a shock would sometimes follow a stimulus. The electrodes were attached to the participant, after which a shock work-up procedure was carried out to set the intensity level of the shock: Stepwise, the electrical current was increased until the participant indicated that the shock was very uncomfortable but not painful. The participant was instructed to give on-line estimates of shock occurrence during each stimulus presentation, and was told that shocks could occur in the next part of the task. After this, the task continued with the acquisition and extinction phase. In the acquisition phase (six presentations of CS<sup>+</sup> and CS<sup>-</sup>), a shock was administered directly after the 8 second CS<sup>+</sup> presentation. Shock duration was 200 ms. Stimuli were presented for 8 s, with a 15 – 25 s variable delay before the next stimulus was presented. In the extinction phase (nine presentations of CS<sup>+</sup> and CS<sup>-</sup>), no shocks were administered. The stimulus-presentation order is presented in Table 3.2.

After the fear conditioning procedure was completed, participants were presented with the fear conditioning questionnaire.

Ethical approval of this study was obtained through the Ethical Committee Psychology of the University of Groningen.

### **Data-analysis**

Per cell of the belief bias design, a single RT score was computed by averaging the median RTs of the two blocks. Only correct responses were included in the calculation of the RT scores. Error scores were computed by summing the errors within each cell of the design. Belief bias scores were computed for the RTs and the errors. The RT belief bias score was computed by subtracting the RTs for the matches from the RTs for the mismatches (cf., Vroling & de Jong, 2010b). Error belief bias scores were calculated in a similar vein.

For the on-line probability estimates, the answer that was given at the end of the stimulus presentation time was defined as the response. The probability estimates for the CS<sup>+</sup> and CS<sup>-</sup> within the acquisition and extinction phase were compared by means of MANOVAs. In each MANOVA (acquisition and extinction), stimulus (CS<sup>+</sup>, CS<sup>-</sup>) was treated as within subject factor and shock expectancies per trial (1 – 6 and 1 – 9 for acquisition and extinction, respectively) were treated as outcome measure.

To be able to correlate acquisition and extinction with belief bias, two partial summary scores were computed for the acquisition and the extinction phase. The shock expectancies for the initial CS<sup>+</sup> and CS<sup>-</sup> presentations were used to check for initial expectancy differences between the stimuli (see above mentioned MANOVAs), but cannot be used to determine the acquisition of fear. The acquisition expectancies have been divided in two parts (Acq1 and Acq2). The first half (Acq1) consisted of the second, third and fourth CS<sup>+</sup> and CS<sup>-</sup>



Table 3.2  
*Stimulus and shock division in both the acquisition and extinction phase (which were summarized in two parts: Acq1 and Acq2 and Ext1 and Ext2 respectively for correlational purposes) of the differential fear conditioning procedure.*

Acquisition		Acq1												Acq2											
trialnumber	1*	Acq1			Acq1									Acq2											
		2	3*	4	5	6	7	8	9	10	11	12													
CS <sup>+</sup> /CS <sup>-</sup>	+	+	-	-	+	-	-	+	-	+	+	-													
Extinction		Ext1												Ext2											
trialnumber	13	Acq2			14*	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30				
		-	+	+	-	-	+	-	-	+	-	-	+	+	-	+	-	+	+	-	+				
CS <sup>+</sup> /CS <sup>-</sup>	-	+	+	+	-	-	-	-	-	+	-	-	+	+	-	+	-	+	+	-	+				

*Note.* When a shock follows the CS<sup>+</sup>, this is indicated by a bold +.  
\* new information presented directly after stimulus presentation

presentations (trials 2, 5, 8 and 4, 6, 7 respectively, which were averaged per CS type), the second half (Acq2) consisted of the fifth, sixth and seventh<sup>7</sup> CS<sup>+</sup> and CS<sup>-</sup> presentations (trials 10, 11, 14 and 9, 12 and 13 respectively, which were averaged per CS type). The extinction phase was divided accordingly: the first half, Ext1, consisted of the trials 16, 19, 22, 23 and 17<sup>8</sup>, 18, 20, 21 for CS<sup>+</sup> and CS<sup>-</sup> respectively (which were averaged per CS type), the second half, Ext2, consisted of the trials 25, 27, 28, 30 and 24, 26, 29 for CS<sup>+</sup> and CS<sup>-</sup>, respectively.

## Results

Due to technical difficulties, one participant had to be excluded. Furthermore, one participant requested to abort the experiment during the differential fear conditioning task. Participants who were incorrect in identifying the CS<sup>+</sup>, or who were less than 50% certain of their identification of the CS<sup>+</sup> (as reported on the fear conditioning questionnaire) were excluded from the analyses, as well as those participants who had not completed the reasoning task according to the instructions (this was determined during the debriefing). In total, 53 participants were included in the final analyses.

### Belief bias task

On average, participants showed both positive belief bias RT scores ( $M = 1962.66$  [in ms],  $SD = 2579.31$ ) and positive belief bias error scores ( $M = 0.43$ ,  $SD = 1.15$ ). These belief bias scores deviated significantly from zero,  $t(52) = 5.5487$ ,  $p < .01$  and  $t(52) = 2.74$ ,  $p < .05$  respectively, indicating that the participants generally showed belief bias which was reflected in both RTs and errors.<sup>9</sup>

### Differential fear conditioning task

Many people (22.06%) were inaccurate in identifying the CS<sup>+</sup>, or were less than 50% certain of their (correct) identification. These people were excluded from the analyses (see above).

---

<sup>7</sup> Note that the seventh CS<sup>+</sup> and CS<sup>-</sup> presentations are actually the beginning of the extinction phase. Yet, because the new information (of both the CS<sup>+</sup> and the CS<sup>-</sup> not being followed by a shock) is only presented directly after the stimulus presentation, the seventh presentation is still part of the acquisition with respect to expectancies.

<sup>8</sup> Note that trial 15 is skipped for the CS<sup>-</sup>. This is done because participants can only develop an understanding of the new rules for shock administration after they have experienced the absence of a shock at CS<sup>+</sup> and then at CS<sup>-</sup> (many participants expected the shock would now be paired with the CS<sup>-</sup>).

<sup>9</sup> The threat related syllogisms that had been included for pilot purposes did not significantly differ from zero, indicating that on average, no threat related belief bias was present. We therefore did not include threat related belief bias in our analyses.

**Acquisition.** On the first trial, there was no difference in UCS expectancy for the CS<sup>+</sup> and the CS<sup>-</sup>,  $F(1,52) = 1.42$ ,  $p = .24$ . Already after the first presentation of the shock (following the CS<sup>+</sup>) the shock expectations differed between the stimuli,  $F(1,52) = 13.13$ ,  $p < .01$ ,  $\eta^2 = .20$  (for the second trial). The difference in UCS expectancy between the CS<sup>+</sup> and the CS<sup>-</sup> continued to enlarge (up until  $\eta^2 = .92$  at the first CS<sup>+</sup> and CS<sup>-</sup> presentation of the extinction phase).

**Extinction.** When the CS<sup>+</sup> was no longer followed by a shock in the extinction phase, the UCS expectancy for the CS<sup>+</sup> decreased from 95.00 to 21.36, while the UCS expectancy on CS<sup>-</sup> trials mildly increased from 6.63 to 16.94. The difference in shock expectation between the CS<sup>+</sup> and the CS<sup>-</sup> however remained significant until the eighth presentation. The difference was no longer significant at the final presentation:  $F_{trial9}(1,52) = 2.86$ ,  $p = .097$ ,  $\eta^2 = .05$ . The differential expectations in the acquisition and in the extinction phase can be seen in Figure 3.1.

### Correlations between belief bias and fear conditioning

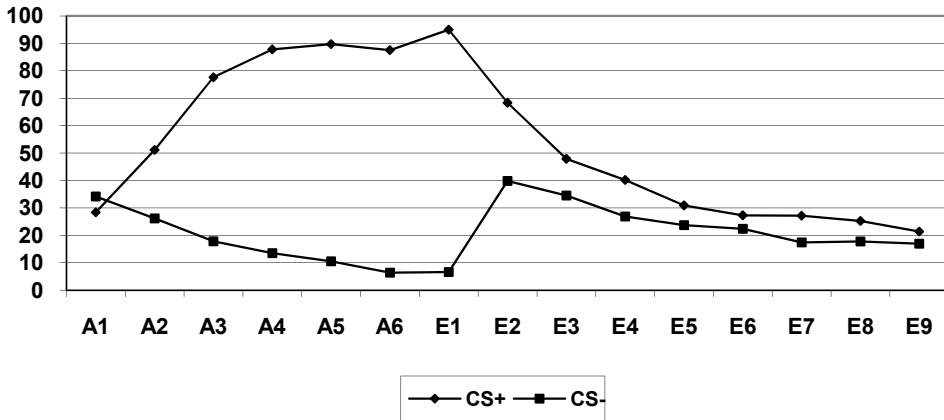
As can be seen in Table 3.3, no meaningful correlations were found between error-based belief bias and differential UCS expectancies during extinction or acquisition. Contrary to our expectations, higher levels of RT-based belief bias were related to lesser differences in UCS expectancy ratings between CS<sup>+</sup> and CS<sup>-</sup> during the first part of extinction,  $r = -.27$ ,  $p = .05$ . In a similar vein, belief bias was also negatively correlated with the first phase of acquisition,  $r = -.29$ ,  $p = .04$ . Thus, participants scoring high on belief bias (measured in RTs) showed *less* differential UCS expectancies during both extinction and acquisition.

Table 3.3. Correlations between belief bias (measured in both errors and RT) and the differential shock expectancy scores for the first and second half of the acquisition and the extinction phase of the fear conditioning procedure of Exp. 1.

	Belief bias errors	Belief bias RT
Acq1: CS <sup>+</sup> - CS <sup>-</sup>	-.12	<b>-.29*</b>
Acq2: CS <sup>+</sup> - CS <sup>-</sup>	.05	-.19
Ext1: CS <sup>+</sup> - CS <sup>-</sup>	-.03	<b>-.27*</b>
Ext2: CS <sup>+</sup> - CS <sup>-</sup>	.01	-.02

\* $p < .10$ , \*\* $p < .05$ , \*\*\* $p < .01$

Figure 3.1. On-line shock expectancies per CS presentation for both the CS<sup>+</sup> and the CS<sup>-</sup> in the acquisition and extinction phase of the fear conditioning experiment (Exp. 1).



## Discussion

The major aim of Experiment 1 was to test whether the extinction of (differential) UCS expectancies varies as a function of people's habitual belief bias. Contrary to expectations, belief bias as indexed by reaction times was negatively correlated to the rate of extinction of differential UCS expectancies.

The pattern of UCS expectancies indicate that both the acquisition and the extinction of differential UCS expectancies were successful. Participants increasingly expected a shock for the CS<sup>+</sup> and no shock for the CS<sup>-</sup> during acquisition and this differential expectancy decreased and finally disappeared during extinction. Participants on average displayed a positive belief bias effect, which indicates that they had more difficulty to respond to syllogisms on which believability and validity did not match. Furthermore, there was considerable variation in the belief bias scores.

It was hypothesized that heightened belief-confirming reasoning (belief bias) would predict delayed extinction. We argued that people with heightened belief bias would have difficulty with incorporating evidence that opposes the previously learned predictive validity of the CS<sup>+</sup>. The present results do not support this view. If anything, the results show the opposite pattern: People with relatively high belief bias showed faster instead of delayed extinction. In a similar vein, participants with relatively strong belief bias also showed relatively slow signal-learning during the acquisition. Thus participants with enhanced belief bias found it relatively difficult to learn the CS-UCS association, and relatively easy to again unlearn this association when the CS<sup>+</sup> was no longer followed by the electrical shock during extinction.

One explanation for these unexpected results could be that we used neutral stimuli as the CS<sup>+</sup>/CS<sup>-</sup> which had no intrinsic relationship with the aversive UCS. The low initial UCS expectancies confirm that participants had no strong a priori beliefs regarding the predictive validity of the CS<sup>+</sup> for shock occurrence. If anything, they seemed to believe that the CS<sup>+</sup> would not be followed by the UCS. Perhaps then it may not be surprising that people with a heightened belief bias showed a relatively strong reluctance to form beliefs regarding the predictive validity of the CS<sup>+</sup> during acquisition, as the CS-UCS pairings ran counter their initial belief (cf. Hamm, Vaitl, & Lang, 1989). In a similar vein, this may explain why people with enhanced belief bias showed less resistance to extinction.

To test whether indeed the unexpected pattern of results can be attributed to the low initial CS-UCS belongingness, we carried out a second experiment, using stimuli with high CS-UCS belongingness.

## Experiment 2

In Experiment 2, we manipulated the a priori UCS belongingness of the CS<sup>+</sup>/CS<sup>-</sup>. In half of the participants we used a CS<sup>+</sup> with low UCS belongingness (i.e., a sunflower) as was the case in Exp. 1, whereas we used a CS<sup>+</sup> with high belongingness (i.e., a cactus) in the other half of the participants. Based on our findings in Exp. 1, we hypothesized to find (again) a negative correlation between habitual belief bias and differential UCS expectancies during extinction when the sunflower was used as the CS<sup>+</sup>, but a positive correlation when the cactus was used as the CS<sup>+</sup>. As a secondary issue we also explored the influence of enhanced belief bias on the acquisition of differential UCS expectancies as a function of a priori CS-UCS belongingness. Since we have now included a between subject factor in our design, our goal was to double the number of participants.

## Method

### Participants

People who were not fluent in Dutch, who suffered from dyslexia, or who had received training on logical reasoning were excluded from entering the study. Participants ( $N = 176$ , 48 male and 128 female) were undergraduate students in psychology (86.4%) and other faculties. The mean age was 20.16 years ( $SD = 3.59$ ). The participating psychology students received either course credits or, if they had already completed their course credit requirements, a small financial reward. The other students always received a small financial reward.

### **Materials and apparatus**

The materials and apparatus were similar to those of Exp. 1, except for the stimuli used in the differential fear conditioning task. The stimuli were a cactus and a sunflower. For half of the participants the cactus served as CS<sup>+</sup> and the sunflower as CS<sup>-</sup> (high belongingness condition) and for half the sunflower served as CS<sup>+</sup> and the cactus as CS<sup>-</sup> (low belongingness condition).

### **Procedure**

The procedure of this experiment was similar to that of Exp. 1, with the following exception: Belongingness condition was randomly determined between participants.

Ethical approval of this study was obtained through the Ethical Committee Psychology of the University of Groningen.

### **Data-analysis**

The data for the high and the low belongingness conditions were analysed separately. The analyses (per belongingness condition) were similar to those of Exp. 1.

### **Results**

Due to technical difficulties, six participants had to be excluded. Also, three participants requested to abort the experiment during the differential fear conditioning task. Furthermore, participants who were incorrect in identifying the CS<sup>+</sup>, or who were less than 50% certain of their identification of the CS<sup>+</sup> were excluded from the analyses, as well as those participants ( $n = 6$ ) who had not completed the reasoning task according to the instructions (this was determined during the debriefing). In total, 151 participants remained in the final analyses, of whom 76 underwent the high belongingness condition and 75 underwent the low belongingness condition.

### **Belief bias task**

On average, participants showed both positive neutral belief bias RT scores ( $M = 1365.23$  [in ms],  $SD = 2546.04$ ) and positive neutral belief bias error scores ( $M = 0.35$ ,  $SD = 1.70$ ). These neutral belief bias scores deviated significantly from zero,  $t(150) = 6.61$ ,  $p < .01$  and  $t(150) = 2.52$ ,  $p < .05$  respectively, indicating that the participants showed belief bias for the neutral themes both on RTs and on errors.<sup>10</sup>

---

<sup>10</sup> Again, we did not find an (on average) belief bias effect for the threat related syllogisms.

**Differential fear conditioning task**

**Low belongingness condition.** Six people (out of the initial 81 in the low belongingness condition, 7.41%) were inaccurate in identifying the CS<sup>+</sup> or were less than 50% certain of their (correct) identification and were excluded from the analyses.

Acquisition: On average, participants initially held higher expectancies for the shock to be followed by the cactus than the sunflower: on the first trial, there was a moderately strong difference in probability estimation between the CS<sup>+</sup> (sunflower,  $M = 29.21$ ) and the CS<sup>-</sup> (cactus,  $M = 46.48$ ),  $F(1,74) = 11.85$ ,  $p < .01$ ,  $\eta^2 = .14$ . This expectation correctly shifted to higher UCS expectancies for the CS<sup>+</sup> as soon as the first shock was administered ( $F[1,74] = 25.55$ ,  $p < .01$ ,  $\eta^2 = .26$  for the second trial). The difference in UCS expectancy between the CS<sup>+</sup> and the CS<sup>-</sup> continued to enlarge (up until  $\eta^2 = .94$  at the first CS<sup>+</sup> and CS<sup>-</sup> presentation of the extinction phase).

Extinction: When the CS<sup>+</sup> (sunflower) was no longer followed by a shock in the extinction phase, the shock expectation for the CS<sup>+</sup> decreased from 95.35 to 20.26, while the CS<sup>-</sup> shock expectancy (cactus) mildly increased from 6.50 to 14.38. The difference in shock expectation between the CS<sup>+</sup> and the CS<sup>-</sup> however remained significant even in the final presentations:  $F(1,74) = 6.41$ ,  $p = .01$ ,  $\eta^2 = .08$ .

**High belongingness condition.** Four participants (out of the initial 80 in the high belongingness condition, 5 %) were inaccurate in identifying the CS<sup>+</sup> or were less than 50% certain of their (correct) identification and were excluded from the analyses.

Acquisition: Similar to the low belongingness condition, participants initially reported higher UCS expectancies for the cactus than for the sunflower: on the first trial, there was a strong difference in probability estimation between the CS<sup>+</sup> (cactus,  $M = 54.12$ ) and the CS<sup>-</sup> (sunflower,  $M = 23.19$ ),  $F(1,75) = 37.85$ ,  $p < .01$ ,  $\eta^2 = .34$ . The difference in UCS expectancy between the CS<sup>+</sup> and the CS<sup>-</sup> continued to enlarge (up until  $\eta^2 = .98$  at the first CS<sup>+</sup> and CS<sup>-</sup> presentation of the extinction phase).

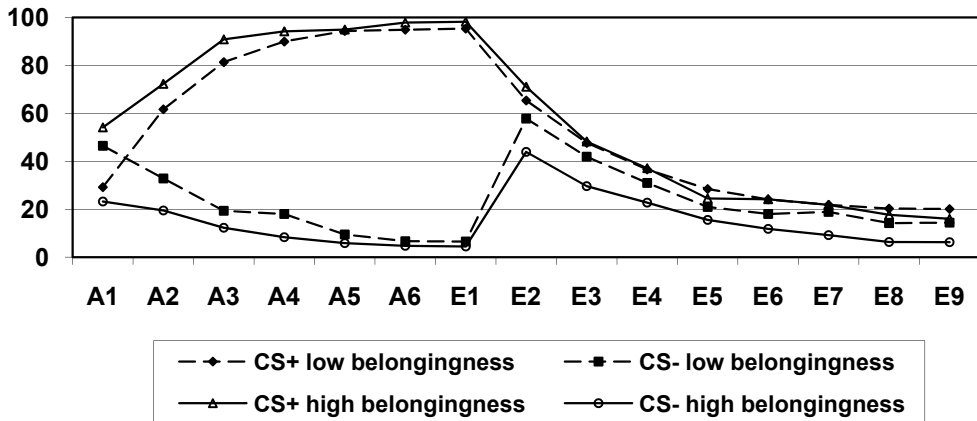
Extinction: When the CS<sup>+</sup> was no longer followed by a shock in the extinction phase, the shock expectation for the CS<sup>+</sup> (cactus) decreased from 98.18 to 17.67, while the CS<sup>-</sup> shock expectancy (sunflower) remained relatively stable from directly prior to the extinction information ( $M = 4.44$ ) to the final presentation within the extinction phase ( $M = 6.28$ ). The difference in shock expectation between the CS<sup>+</sup> and the CS<sup>-</sup> however remained significant even in the final presentations:  $F(1,75) = 21.19$ ,  $p < .01$ ,  $\eta^2 = .22$ .

The differential shock expectancies in the low and the high belongingness condition for the acquisition and the extinction phase are shown in Figure 3.2.

### Correlations between belief bias and low belongingness fear conditioning

No significant correlations were found between belief bias and differential shock expectancies in the low belongingness condition.

Figure 3.2. On-line shock expectancies for the high belongingness (cactus as CS<sup>+</sup>, sunflower as CS<sup>-</sup>) and the low belongingness condition (sunflower as CS<sup>+</sup>, cactus as CS<sup>-</sup>) per CS presentation for both the CS<sup>+</sup> and the CS<sup>-</sup> in the acquisition and extinction phase of the fear conditioning experiment (Exp. 2).



### Correlations between belief bias and high belongingness fear conditioning

In line with our hypothesis, higher belief bias error scores were related to larger differential expectancies in both phases of the extinction,  $r_{\text{Ext1, BBerror}} = .23, p < .05$  and  $r_{\text{Ext2, BBerror}} = .24, p < .05$ , respectively. Participants high on belief bias (measured in errors) indeed showed delayed extinction. This pattern was absent for the RT-based belief bias scores. In addition, no significant correlations were found between belief bias and differential UCS expectancies during the acquisition phases. The correlations for the high belongingness condition are presented in Table 3.4.



Table 3.4

*Correlations between belief bias (measured in both errors and RT) and the differential shock expectancy scores for the first and second half of the acquisition and the extinction phase of the fear conditioning procedure of the high belongingness condition of Experiment 2.*

	Belief bias errors	Belief bias RT
Acq1: CS <sup>+</sup> - CS <sup>-</sup>	-.09	.04
Acq2: CS <sup>+</sup> - CS <sup>-</sup>	-.16	.04
Ext1: CS <sup>+</sup> - CS <sup>-</sup>	<b>.23*</b>	-.06
Ext2: CS <sup>+</sup> - CS <sup>-</sup>	<b>.24*</b>	-.06

\* $p < .10$ , \* $p < .05$ , \*\* $p < .01$

## Discussion

The major aim of Exp. 2 was to test whether in case of high prior CS-UCS belongingness, belief bias would be related to delayed extinction. This is indeed what we found: People with a higher tendency to engage in belief-biased reasoning fallacies showed slower extinction of UCS expectancies on CS<sup>+</sup> trials when the CS<sup>+</sup> was a picture of a cactus (viz. the high belongingness condition).

Supporting the validity of stimulus selection, participants in both conditions initially showed higher UCS expectancies for the cactus than for the sunflower. Yet, regardless of belongingness, the acquisition and extinction of fear was successful in both conditions, although shock expectancies remained slightly higher for the CS<sup>+</sup> than the CS<sup>-</sup> at the end of the extinction phase (again in both conditions).

Our predictions were only confirmed for error-based belief bias and not for reaction time based belief bias. People who have a tendency for biased reasoning but who manage to correct themselves, the critical and therefore careful reasoners, do not show delayed extinction in this experiment. The time-frame within which participants were asked to rate their UCS-expectancy was 8 seconds. An 8-second time-frame facilitates the correction of initial belief-based UCS expectancies in critical reasoners. In daily-life however, decision-time is usually much smaller and situations are more ambiguous, which will likely leave too little time and attention for the critical reasoner to signal his initial reasoning error and to correct it. Following this, also reaction time based belief bias will be predictive of delayed extinction when the situation would require an immediate indication of participants' UCS expectancy.

In apparent contrast to the results of Exp. 1, there was no relationship between belief bias and the acquisition of UCS expectancies for the high belongingness condition. However, in Exp. 1, the relationship between belief bias and acquisition was limited to the first phase of the acquisition. The

pattern of acquisition of Exp. 2 differs from Exp. 1 on exactly this phase: initial differential expectancies are larger in Exp. 2 and therefore acquisition is somewhat more rapid, therefore probably leaving less room for belief bias to affect expectancies. The same holds for the acquisition phase of the low belongingness condition. Also, the results from the fear conditioning questionnaire indicate that detecting the CS<sup>+</sup>-UCS relationship was easier in Exp. 2 than Exp. 1, which will likely further decrease variability in Exp. 2.

In apparent contrast to Exp. 1, there was no relationship between the strength of belief bias and the rate of extinction in the low belongingness condition. However, one important difference between both experiments is that in Exp. 1 we used neutral CS<sup>+</sup>/CS<sup>-</sup> stimuli without any belongingness with the UCS, whereas in the low belongingness condition of Exp. 2 the CS<sup>-</sup> had a strong belongingness with the UCS. Because the CS-UCS contingency ran counter the strong a priori belongingness between cactus-and-shock and sunflower-and-no-shock, participants were generally very quick in learning that the sunflower was no longer followed by a shock during extinction. Clearly, this reduced the variability in participants' rate of extinction, thereby reducing the room for enhanced belief bias to increase further the rate of extinction. Thus the absence of a relationship between belief bias and both rate of acquisition and extinction in the low belongingness condition can be explained by the reduced variability of participants' pattern of UCS expectancies in Exp. 2.

## General Discussion

Classical conditioning theory is still one of the most influential theories on phobic anxiety. Contemporary learning theories have added the notion of individual variability in conditionability. Various studies have made clear that anxiety disorder patients show stronger conditionability (on either extinction or both acquisition and extinction; e.g., Michael et al., 2007; Orr et al., 2000) yet there have been only few studies attempting to predict (or explain) this variability in conditionability. These few studies have mainly focussed on personality traits (e.g., Davey & Matchett, 1994; Orr et al., 2000; Otto et al., 2007; Pineles, Vogt, & Orr, 2009). Although it has become clear in recent years that information processing biases are influential in the maintenance and sometimes also the development of anxiety disorders (e.g., Mathews & MacLeod, 2005), no studies have been reported that attempt to predict individual differences in conditionability by individual differences in such cognitive biases. The present study provided evidence to suggest that differential information-processing biases might nevertheless be considered likely candidates in the unravelling of individual differences in resistance to extinction of UCS

expectancies and, more generally, potentially also in the unravelling of differences in resistance to the extinction of fear.

The goal of the present experiments was to investigate whether individual differences in habitual belief bias may be one of the factors that can help explain individual differences in the rate of fear extinction. These differences have already been linked to anxiety disorders: Previous work showed that for instance panic disorder patients are characterized by a slower extinction of fear than non-anxious controls (Michael et al., 2007), yet little is known on what constitutes this delayed extinction. The present findings show that enhanced belief bias may be one of the factors that may help explaining such delayed extinction of UCS expectancies. In addition, the present pattern of findings clearly show that the influence of enhanced belief bias critically depends on the nature of the stimuli used to condition fear: If the stimuli are UCS-relevant, belief bias was found to be related to delays in the extinction of UCS expectancies (Exp. 2, high belongingness condition). If, however, the stimuli involved in the fear conditioning show no *a priori* meaningful relationship with the UCS, enhanced belief bias may serve as a protective factor (Exp. 1). Given that most anxiety disorders usually concern fear-relevant stimuli with high *a priori* UCS belongingness (e.g., Seligman, 1971), belief bias will most often serve to maintain anxiety disorders.

The finding of belief bias as predictor for delayed extinction fits in with and adds to earlier findings in the field of belief bias and psychopathology: de Jong, Weertman and colleagues (1997) found that spider phobic patients are characterized by a generally enhanced belief bias and they concluded that perhaps belief bias may have contributed to the development of the phobia. Later studies failed to find a relationship between belief bias and psychopathologic symptoms in general samples (Smeets & de Jong, 2005; Vroling & de Jong, 2010b), yet it was noted that no experimental control was exerted over the learning experiences that may contribute to the development of anxiety symptoms. The present findings, which were found by exerting experimental control over the learning experiences, do support such relationship and fit in with the notion that belief bias may contribute to the development of a specific phobia by contributing to a continued UCS expectancy for high UCS belongingness stimuli (cf. Merckelbach, de Jong, Muris, & van den Hout, 1996). Also, the present findings are in line with the recent suggestion by Mitchell, De Houwer and Lovibond (2009) that reasoning is critically involved in classically conditioned fear learning.

It would be interesting for future research to bring belief bias under experimental control. Given the present findings, it is expected that experimentally induced low belief biased reasoning will facilitate the extinction of differential UCS expectancies when there is a strong *a priori* belongingness

between the CS<sup>+</sup> and the UCS, whereas experimentally induced high levels of belief bias will result in a delayed extinction of UCS expectancies. If so, it would be worth looking into the possibilities to implement a belief bias modification training as a preventive measure (cf. e.g., Holmes, Lang, & Shah, 2009; Schartau, Dalgleish, & Dunn, 2009).

To summarize, the present findings support the view that a habitual belief-confirming reasoning strategy (viz. belief bias) may be involved in delayed extinction of UCS expectancies and can therefore be seen as one of the factors that may contribute to the consolidation of dysfunctional convictions through which anxiety disorders may develop/maintain. The present study underlines the importance of investigating the possible contribution of reasoning biases in aversive conditioning, and may help explaining individual differences in the extinction of fear.

